

文章编号: 1002-1175(2008)05-0665-06

# 实时校正模型的抗差递推算法<sup>\*</sup>

赵 超<sup>†</sup> 洪华生 张珞平

(近海海洋环境科学国家重点实验室, 厦门大学环科中心, 厦门 361005)

(2007年9月21日收稿; 2008年1月14日收修改稿)

Zhao C, Hong HS, Zhang LP. Robust recursive algorithm for real-time correction model. *Journal of the Graduate School of the Chinese Academy of Sciences*, 2008, 25(5): 665~670

**摘 要** 利用遥测系统实时监测水情资料, 由于遥测系统自身的原因以及水文要素测量的具体要求, 数据常常携带异常误差. 采用有异常误差的实测流量资料对实时校正模型进行参数辨识, 要求算法既能抵御异常误差的影响, 又具有较强的实时跟踪能力, 以适应实时洪水预报的要求. 在递推最小二乘算法的基础上, 引入抗差理论, 削弱异常值对参数估计的影响; 引入遗忘因子, 实时跟踪模型时变参数的变化. 计算实例表明, 带有遗忘因子的抗差递推最小二乘算法对异常误差不敏感, 又具有较强的实时跟踪能力.

**关键词** 校正模型, 抗差理论, 递推算法, 遗忘因子

**中图分类号** P338

当水文预报模型采用概念性水文模型, 如新安江模型时, 实时校正模型可采用误差自回归(AR)模型. 若AR模型误差服从正态分布, 递推最小二乘(RLS)算法可获得参数最优估值.

随着科学技术的发展, 实时水情资料大都通过遥测系统自动监测获得. 由于遥测系统自身以及水文要素测量的具体要求, 使得水情数据中不可避免地存在一些未知分布的异常误差. 例如, 对于水库流域来说, 流域实测流量不能直接获得. 多是通过实时监测库水位的变化和出库水量, 采用水量平衡方程获取流域的实测流量. 当在大洪水期间测量库水位时, 易受大风影响, 库水位常存在误差, 再换算成实测流量值, 常出现异常误差. 此时, 采用RLS算法获得的参数估值受异常误差影响剧烈, 偏离真值.

本文将抗差估计引入实时校正模型参数估计中, 利用抗差系统具有的抗差能力<sup>[1~3]</sup>, 减小估计方法受异常误差的影响, 推导出带有遗忘因子的抗差递推最小二乘(RRLS)算法, 估计有异常误差下的最佳参数, 从而提高实时洪水预报精度.

## 1 抗差递推最小二乘算法

抗差估计理论是一门十分年轻的分支学科, 抗差估计实际是在异常值不可避免的情况下, 选择适当的估计方法使未知量估值尽可能减免异常值的影响, 得出正常模式下的最佳估值.

周江文等<sup>[4]</sup>将抗差估计理论与最小二乘算法相结合, 得到抗差最小二乘算法. 本文进一步将抗差最小二乘算法与有遗忘因子的递推算法相结合, 获得有抵御异常值影响的时变参数动态估计方法.

### 1.1 问题的描述

<sup>\*</sup> 国家自然科学基金项目(40701178)资助

<sup>†</sup>E-mail: zhaocfang@sina.com

误差自回归( auto regression, AR) 模型, 是利用径流量实测值和水文模型预报值之间的误差序列的自相关性, 利用过去的误差预报未来的误差. 一  $n$  阶 AR 模型可表示为:

$$y(t) = \mathbf{X}_t^T \boldsymbol{\theta} + \varepsilon(t), \quad (t = 1, 2, \dots) \tag{1}$$

其中,  $y(t)$  为  $t$  时刻实测流量与水文模型预报流量的偏差. 当实测流量存在异常误差, 就转化为  $y$  存在异常误差;  $\varepsilon$  为残差;  $\mathbf{X}_t = [y(t-1), y(t-2), \dots, y(t-n)]^T$ ;  $\boldsymbol{\theta}$  为参数向量.

综合考虑前  $t$  次的测量结果, 可得

$$\mathbf{Y}_t = \boldsymbol{\Phi} \boldsymbol{\theta} + \mathbf{E}_t. \tag{2}$$

式中,  $\mathbf{Y}_t = \begin{bmatrix} y(1) \\ \vdots \\ y(t-1) \\ y(t) \end{bmatrix}$ ,  $\boldsymbol{\Phi} = \begin{bmatrix} \mathbf{X}_1^T \\ \vdots \\ \mathbf{X}_{t-1}^T \\ \mathbf{X}_t^T \end{bmatrix}$ ,  $\mathbf{E}_t = \begin{bmatrix} \varepsilon(1) \\ \vdots \\ \varepsilon(t-1) \\ \varepsilon(t) \end{bmatrix}$ .

为了减小旧数据的信息量, 强调新信息, 对旧的残差逐次加遗忘因子  $\beta$  ( $\beta \leq 1$ ), 则残差向量为

$$\mathbf{E}_t = [\beta^{t-1} \varepsilon(1), \beta^{t-2} \varepsilon(2), \dots, \beta \varepsilon(t-1), \varepsilon(t)]^T. \tag{3}$$

因为,

$$\beta \varepsilon = \beta y - \beta \mathbf{X}^T \boldsymbol{\theta}. \tag{4}$$

则(2)式中各矩阵变化为

$$\mathbf{Y} = \begin{bmatrix} \beta^{t-1} y(1) \\ \vdots \\ \beta y(t-1) \\ y(t) \end{bmatrix}, \quad \boldsymbol{\Phi} = \begin{bmatrix} \beta^{t-1} \mathbf{X}_1^T \\ \vdots \\ \beta \mathbf{X}_{t-1}^T \\ \mathbf{X}_t^T \end{bmatrix}, \quad \mathbf{E}_t = \begin{bmatrix} \beta^{t-1} \varepsilon(1) \\ \vdots \\ \beta \varepsilon(t-1) \\ \varepsilon(t) \end{bmatrix}.$$

当  $y$  中不可避免存在异常误差时,  $\varepsilon$  将不再服从正态分布, 而成为一个正态混合分布, 可表示为<sup>[4]</sup>

$$p(\varepsilon) = (1 - \delta) N(\varepsilon | 0, \sigma^2) + \delta h(\varepsilon), \quad 0 \leq \delta \leq 1, \tag{5}$$

式中,  $h(\cdot)$  为污染分布,  $\delta$  为污染率, 表示异常值的数据在整个数据中占的比例.

### 1.2 算法的推导

根据抗差最小二乘算法的估计准则<sup>[4]</sup> 可得

$$\frac{\partial J}{\partial \boldsymbol{\theta}} = (\boldsymbol{\Phi}_t^T \mathbf{W}_t \boldsymbol{\Phi}_t \boldsymbol{\theta} - \boldsymbol{\Phi}_t^T \mathbf{W}_t \mathbf{Y}_t) + (\boldsymbol{\Phi}_t^T \mathbf{W}_t \boldsymbol{\Phi}_t \boldsymbol{\theta} - \boldsymbol{\Phi}_t^T \mathbf{W}_t \mathbf{Y}_t)^T = 0, \tag{6}$$

式中,  $\mathbf{W}_t = \begin{bmatrix} w(1) & & 0 \\ & w(2) & \\ & & \dots \\ 0 & & & w(t) \end{bmatrix}$ ,  $w(t)$  为抗差等价权, 表示  $y(t)$  的权重, 即  $y(t)$  的可信度. 得

$$\boldsymbol{\theta}(t) = \mathbf{P}_t \boldsymbol{\Phi}_t^T \mathbf{W}_t \mathbf{Y}_t, \tag{7}$$

其中,  $\mathbf{P}_t = (\boldsymbol{\Phi}_t^T \mathbf{W}_t \boldsymbol{\Phi}_t)^{-1}$ . (8)

当计入第  $t+1$  时刻的数值时, 考虑遗忘因子  $\beta$  的残差向量为

$$\mathbf{E}_{t+1} = \begin{bmatrix} \beta^t \varepsilon(t) \\ \vdots \\ \beta \varepsilon(t) \\ \varepsilon(t+1) \end{bmatrix} = \begin{bmatrix} \beta \mathbf{E}_t \\ \varepsilon(t+1) \end{bmatrix}. \tag{9}$$

综合考虑所有的结果, 有  $\mathbf{Y}_{t+1} = \boldsymbol{\Phi}_{t+1} \boldsymbol{\theta} + \mathbf{E}_{t+1}$ , (10)

式中,  $\mathbf{Y}_{t+1} = \begin{bmatrix} \beta^t y(1) \\ \vdots \\ \beta y(t) \\ y(t+1) \end{bmatrix}$ ,  $\boldsymbol{\Phi}_{t+1} = \begin{bmatrix} \beta^t \mathbf{X}_1^T \\ \vdots \\ \beta \mathbf{X}_t^T \\ \mathbf{X}_{t+1}^T \end{bmatrix} = \begin{bmatrix} \beta \boldsymbol{\Phi}_t \\ \mathbf{X}_{t+1}^T \end{bmatrix}$ ,

可获得与式(7)、式(8)相似的结果:

$$\theta(t+1) = \mathbf{P}_{t+1} \Phi_{t+1}^T \mathbf{W}_{t+1} \mathbf{Y}_{t+1}, \quad (11)$$

$$\mathbf{P}_{t+1} = (\Phi_{t+1}^T \mathbf{W}_{t+1} \Phi_{t+1})^{-1}, \quad (12)$$

$$\text{式中, } \mathbf{W}_{t+1} = \begin{bmatrix} w(1) & & & 0 \\ & \dots & & \\ & & w(t) & \\ 0 & & & w(t+1) \end{bmatrix} = \begin{bmatrix} \mathbf{W}_t & 0 \\ 0 & w(t+1) \end{bmatrix}.$$

由式(12)可得

$$\begin{aligned} \mathbf{P}_{t+1}^{-1} &= \begin{bmatrix} \beta \Phi \\ \mathbf{X}_{t+1}^T \end{bmatrix}^T \begin{bmatrix} \mathbf{W}_t & 0 \\ 0 & w(t+1) \end{bmatrix} \begin{bmatrix} \beta \Phi \\ \mathbf{X}_{t+1}^T \end{bmatrix} = \beta^2 \Phi_t^T \mathbf{W}_t \Phi_t + \mathbf{X}_{t+1}^T w(t+1) \mathbf{X}_{t+1} \\ &= \beta^2 \mathbf{P}_t^{-1} + \mathbf{X}_{t+1}^T w(t+1) \mathbf{X}_{t+1}. \end{aligned} \quad (13)$$

对式(13)利用矩阵求逆引理<sup>[5]</sup>, 可得

$$\mathbf{P}_{t+1} = \frac{1}{\beta^2} \{ \mathbf{P}_t - \mathbf{P}_t \mathbf{X}_{t+1} [\beta^2 w^{-1}(t+1) + \mathbf{X}_{t+1}^T \mathbf{P}_t \mathbf{X}_{t+1}]^{-1} \mathbf{X}_{t+1}^T \mathbf{P}_t \}, \quad (14)$$

$$\begin{aligned} \theta(t+1) &= \mathbf{P}_{t+1} \begin{bmatrix} \beta \Phi \\ \mathbf{X}_{t+1}^T \end{bmatrix}^T \begin{bmatrix} \mathbf{W}_t & 0 \\ 0 & w(t+1) \end{bmatrix} \begin{bmatrix} \beta \mathbf{Y}_t \\ y(t+1) \end{bmatrix} \\ &= \left\{ \frac{1}{\beta^2} \mathbf{P}_t - \frac{1}{\beta^2} \mathbf{P}_t \mathbf{X}_{t+1} [\beta^2 w^{-1}(t+1) + \mathbf{X}_{t+1}^T \mathbf{P}_t \mathbf{X}_{t+1}]^{-1} \mathbf{X}_{t+1}^T \mathbf{P}_t \right\} \\ &\quad \cdot [\beta^2 \Phi_t^T \mathbf{W}_t \mathbf{Y}_t + \mathbf{X}_{t+1}^T w(t+1) y(t+1)] \\ &= \mathbf{P}_t \mathbf{X}_t^T \mathbf{W}_t \mathbf{Y}_t - \mathbf{P}_t \mathbf{X}_{t+1} [\beta^2 w^{-1}(t+1) + \mathbf{X}_{t+1}^T \mathbf{P}_t \mathbf{X}_{t+1}]^{-1} \cdot \mathbf{X}_{t+1}^T \mathbf{P}_t \Phi_t \mathbf{W}_t \mathbf{Y}_t + \frac{1}{\beta^2} \mathbf{P}_t \mathbf{X}_{t+1} w(t+1) y(t+1) \\ &\quad - \frac{1}{\beta^2} \mathbf{P}_t \mathbf{X}_{t+1} [\beta^2 w^{-1}(t+1) + \mathbf{X}_{t+1}^T \mathbf{P}_t \mathbf{X}_{t+1}]^{-1} \cdot \mathbf{X}_{t+1}^T \mathbf{P}_t \mathbf{X}_{t+1} w(t+1) y(t+1). \end{aligned} \quad (15)$$

(15) 式中后 2 项为

$$\begin{aligned} &\frac{1}{\beta^2} \mathbf{P}_t \mathbf{X}_{t+1} w(t+1) y(t+1) - \frac{1}{\beta^2} \mathbf{P}_t \mathbf{X}_{t+1} [\beta^2 w^{-1}(t+1) + \mathbf{X}_{t+1}^T \mathbf{P}_t \mathbf{X}_{t+1}]^{-1} \cdot \mathbf{X}_{t+1}^T \mathbf{P}_t \mathbf{X}_{t+1} w(t+1) y(t+1) \\ &= \frac{1}{\beta^2} \mathbf{P}_t \mathbf{X}_{t+1} w(t+1) y(t+1) - \frac{1}{\beta^2} \mathbf{P}_t \mathbf{X}_{t+1} w(t+1) y(t+1) \\ &\quad + \mathbf{P}_t \mathbf{X}_{t+1} [\beta^2 w^{-1}(t+1) + \mathbf{X}_{t+1}^T \mathbf{P}_t \mathbf{X}_{t+1}]^{-1} y(t+1) \\ &= \mathbf{P}_t \mathbf{X}_{t+1} [\beta^2 w^{-1}(t+1) + \mathbf{X}_{t+1}^T \mathbf{P}_t \mathbf{X}_{t+1}]^{-1} y(t+1). \end{aligned} \quad (16)$$

将式(7)、(16)代入式(15)得

$$\theta(t+1) = \theta(t) + w(t+1) \mathbf{P}_t \mathbf{X}_{t+1} [\beta^2 + w(t+1) \mathbf{X}_{t+1}^T \mathbf{P}_t \mathbf{X}_{t+1}]^{-1} [y(t+1) - \mathbf{X}_{t+1}^T \theta(t)]. \quad (17)$$

$$\mathbf{P}_{t+1} = \frac{1}{\beta^2} \left\{ \mathbf{I} - w(t+1) \mathbf{P}_t \mathbf{X}_{t+1} [\beta^2 + w(t+1) \mathbf{X}_{t+1}^T \mathbf{P}_t \mathbf{X}_{t+1}]^{-1} \mathbf{X}_{t+1}^T \right\} \mathbf{P}_t. \quad (18)$$

式中,  $\mathbf{I}$  为单位矩阵.

式(17)和式(18)构成了完整的带有遗忘因子的抗差递推最小二乘算法. 遗忘因子  $\beta$  降低了旧数据的信息量, 增加了新数据对估计的影响, 以使估计实时跟踪时变数据的变化.  $\beta^2$  一般采用试错法获得, 对于水文变量的预报问题,  $\beta^2$  取值一般在 0.9 到 1.0 之间.  $w(t)$  表征  $y(t)$  的可信度, 在洪水预报实时校正时, 本文采用的抗差等价权函数<sup>[6]</sup>

$$w(t) = \begin{cases} w(t) & | \varepsilon(t) | \leq 1.5\phi \\ w(t) 1.5\phi / | \varepsilon(t) | & 1.5\phi < | \varepsilon(t) | \leq 2.5\phi, \\ 0 & | \varepsilon(t) | > 2.5\phi \end{cases} \quad (19)$$

式中,  $\phi$  为加权残差均方差.

当残差绝对值较小时( $\leq 1.5\phi$ ), 权重保持不变, 表示  $y$  可信; 当残差绝对值较大时( $> 2.5\phi$ ), 权重为 0, 表示  $y$  异常, 干扰严重, 将其淘汰; 当残差绝对值位于两者之间时, 采取降权处理, 减小其对参数估计的影响. 由于 RRLS 算法中  $w(\bullet)$  是残差的函数, 所以需要采用迭代计算, 过程如下:

- (1) 首先假设  $w(t+1) = 1$ , 根据式(17)、(18) 计算  $\theta^{(1)}(t+1)$ ,  $P_{t+1}$ , 计算出  $\varepsilon^{(1)}(t+1)$ ;
- (2) 按式(19) 计算  $t+1$  时刻的等价权;
- (3) 重复(1)、(2), 迭代权重, 直到  $|\theta^{(k)}(t+1) - \theta^{(k+1)}(t+1)| < \mu$ , 迭代停止,  $\theta(t+1) = \theta^{(k+1)}(t+1)$ . 迭代停止时, 若  $w(t+1) = 0$ , 可判断  $y(t+1)$  为异常值, 根据式(17) 计算出的  $\theta(t+1)$  与  $\theta(t)$  相同, 抵御了异常值对参数估计的影响.

RRLS 算法需要初始条件  $\theta(0)$ ,  $P(0)$ . 考虑到实时洪水与历史洪水存在相似性, 实时洪水是历史洪水平均条件下的振荡, 所以本文采用历史洪水的平均参数作为  $\theta(0)$ , 相应的  $P$  值作为  $P(0)$ .

## 2 计算实例

本文采用的资料为闽江七里街流域 1988~ 1998 年间人工监测洪水水情资料, 由于是人工监测, 所以此类资料中不含异常误差. 其中 1988~ 1996 年的 36 场视为历史洪水, 用来试算,  $\beta^2 = 0.96$ ,  $\theta(0) = 0.9848$ . 1997~ 1998 年的 7 场洪水视作实时洪水. 水文预报模型采用新安江模型, 实时校正模型采用一阶 AR 模型. 实时校正的预见期为 3 h.  $A_1$  为 RLS 的算法结果,  $A_2$  为 RRLS 的算法结果.

### 2.1 无异常值情况

首先将 RRLS 算法与 RLS 算法分别运用于无异常值情况. 很明显, 此时 RLS 算法可得到无偏最优估值. 本文将此时 RLS 算法获得的参数值作为近似真值的参数值, 记为  $\theta_0$ . 表 1 显示了无异常值时, RRLS 算法获得的参数  $\theta_r$  相对于  $\theta_0$  的离散程度. 用 2 种算法获得的参数估值进行实时修正, 洪水过程线确定性系数的比较见表 2.

表 1 无异常值时 RRLS 算法参数估值的离散程度

Flood	V	Flood	V
980608	0.0217	970701	0.0009
980509	0.0048	970618	0.0041
980301	0.0048	970605	0.0005
980215	0.0019		

注:  $V = \left\{ \frac{1}{m-1} \sum_{i=1}^m [\theta_{A_2}(i) - \theta_0(i)]^2 \right\}^{0.5}$ ,  $m$  为时段长.

表 2 无异常值时两种方法预报精度

洪水	Dc(A <sub>1</sub> )	Dc(A <sub>2</sub> )	洪水	Dc(A <sub>1</sub> )	Dc(A <sub>2</sub> )
980608	0.966	0.962	970701	0.976	0.976
980509	0.958	0.942	970618	0.969	0.978
980301	0.976	0.978	970605	0.889	0.905
980215	0.965	0.959			

注:  $Dc = 1 - \frac{\sum (Q_i - \bar{Q})^2}{\sum (Q_i - \bar{Q})^2}$ ,  $Q$  为实测流量;  $Q_i$  为预报流量;  $\bar{Q}$  为实测流量均值.

表 1 和表 2 的结果很明显, 当实测流量不存在异常误差时, RRLS 算法的结果与 RLS 算法的结果非常接近. 说明此时抗差递推算法获得的参数估值非常接近最优值.

### 2.2 有异常值情况

对前面的无异常值人工监测数据, 添加异常误差, 获得有异常值的数据. 这样异常值的分布和数目可以人工控制, 从而能清楚地分析出估计方法抵抗异常值的效果.

采用的异常误差生成模式为

$$e_i = \begin{cases} (r - 0.5) \bar{Q} p & i = \text{int}(i/L) L \\ 0 & i \neq \text{int}(i/L) L \end{cases} \quad (20)$$

式中,  $r$  为随机数;  $\bar{Q}$  为实测流量的均值, 考虑添加的异常误差与洪水量级有关;  $p$  为常数, 可控制异常误差的大小;  $L$  为异常值产生的频率间隔, 可控制异常误差产生的频率. 对  $p$  和  $L$  的随机选取几个值, 来分析不同量级、不同频率的异常误差对估计方法的影响. 本文选取以下 4 种组合: (1)  $p = 3$ ,  $L = 10$ ; (2)

$p = 3, L = 15; (3) p = 5, L = 10; (4) p = 5, L = 15.$

表 3 显示了 2 种算法在 4 种组合下获得的参数估值相对于  $\theta_o$  的离散程度; 采用 2 种算法获得的参数估值进行洪水预报实时校正, 洪水校正过程线确定性系数见表 4.

表 3 4 种组合下 2 种算法离散程度

Flood	$p = 3, L = 10$		$p = 3, L = 15$		$p = 5, L = 10$		$p = 5, L = 15$	
	$V_{A_1}$	$V_{A_2}$	$V_{A_1}$	$V_{A_2}$	$V_{A_1}$	$V_{A_2}$	$V_{A_1}$	$V_{A_2}$
980608	0.2294	0.0116	0.2294	0.0116	0.2856	0.0141	0.2765	0.0137
980509	0.0510	0.0031	0.0510	0.0031	0.1140	0.0020	0.1177	0.0020
980301	0.2910	0.0035	0.2910	0.0035	0.3508	0.0087	0.3168	0.0023
980215	0.2331	0.0020	0.2331	0.0020	0.2982	0.0021	0.2700	0.0010
970701	0.1939	0.0042	0.1939	0.0042	0.2742	0.0042	0.2683	0.0049
970618	0.2249	0.0043	0.2249	0.0043	0.2990	0.0048	0.2495	0.0012
970605	0.0195	0.0003	0.0195	0.0003	0.0496	0.0003	0.0371	0.0006

注:  $V_{A_1} = \left\{ \frac{1}{m-1} \sum_{i=1}^m [\theta_{A_1}(i) - [\theta_o(i)]]^2 \right\}^{0.5}, V_{A_2} = \left\{ \frac{1}{m-1} \sum_{i=1}^m [\theta_{A_2}(i) - \theta_o(i)]^2 \right\}^{0.5}$

表 4 4 种组合下 2 种算法校正精度

Flood	$p = 3, L = 10$		$p = 3, L = 15$		$p = 5, L = 10$		$p = 5, L = 15$	
	Dc(A <sub>1</sub> )	Dc(A <sub>2</sub> )	Dc(A <sub>1</sub> )	Dc(A <sub>2</sub> )	Dc(A <sub>1</sub> )	Dc(A <sub>2</sub> )	Dc(A <sub>1</sub> )	Dc(A <sub>2</sub> )
980608	0.893	0.964	0.893	0.964	0.841	0.961	0.802	0.966
980509	0.917	0.959	0.917	0.959	0.861	0.958	0.910	0.957
980301	0.902	0.971	0.902	0.971	0.800	0.972	0.833	0.974
980215	0.736	0.958	0.736	0.958	0.456	0.946	0.642	0.953
970701	0.857	0.975	0.857	0.975	0.736	0.971	0.793	0.973
970618	0.897	0.968	0.897	0.968	0.798	0.966	0.861	0.966
970605	0.782	0.900	0.782	0.900	0.633	0.887	0.756	0.886

分析表 3 可以发现, 对于选取的不同  $p$  和  $L$ , RLS 算法由于对异常值比较敏感, 参数估值偏离真值较远; 同时随着异常误差量级的增大( $p$  增大) 和间隔频率的缩短( $L$  减小), 估计结果偏离真值越严重. 而 RRLS 算法对异常值赋 0 权, 抵御了异常值对参数估计的影响, 4 种组合下均能获得较稳定且接近于  $\theta_o$  的参数估值.

分析表 4 同样发现, RLS 算法的校正精度明显受异常值的影响, 随着异常误差量级的增大( $p$  增大) 和间隔频率的缩短( $L$  减小), 预报精度降低. 而 RRLS 算法在不同  $p$  和  $L$  的情况下, 均可获得较稳定且精度较高的校正精度.

由于计算结果类似, 本文仅选取洪水 980608, 显示了 2 种算法在 4 种组合下的估值与  $\theta_o$  的比较(见图 1 和图 2).

分析图 1 和图 2, 当有异常值存在时, RRLS 算法的参数估值与  $\theta_o$  差别不大, 较好地跟踪了时变参数的变化, 而 RLS 算法的参数估值已远远偏离了  $\theta_o$ , 且随着异常误差量级的增大( $p$  增大) 和间隔频率的缩短( $L$  减小), 偏离的越严重. 说明 RRLS 算法对异常值不敏感, 而 RLS 非常敏感.

3 结论

本文推导出的带有遗忘因子的抗差递推最小二乘算法, 既能抵御异常值对参数估计的影响, 又能跟踪时变参数的动态变化. 将此运用到洪水预报的实时校正模型当中, 计算结果表明, 无论实测资料是否

存在异常值,RRLS 算法都能获得高精度的实时校正结果. 这为提高实时洪水预报精度提供了一种新的方法.

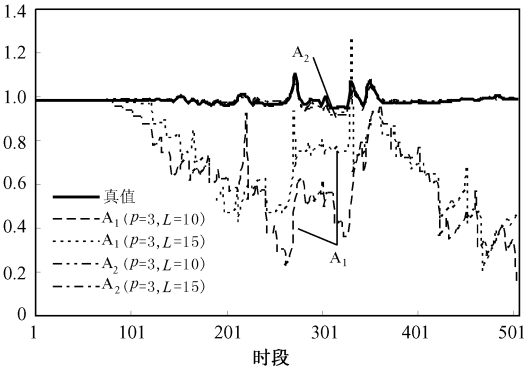


图 1 洪水 960608 两种算法的参数估值( $p=3$ )

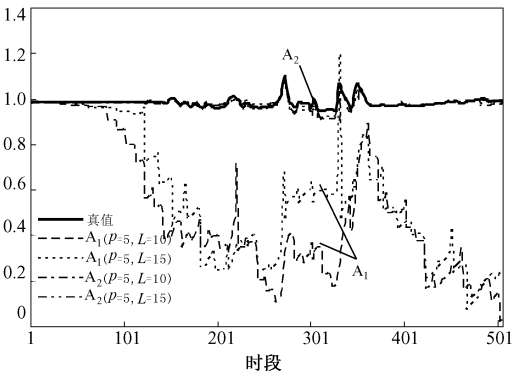


图 2 洪水 960608 两种算法的参数估值( $p=5$ )

今后的研究工作一方面要进一步完善抗差递推算法,如抗差权函数的选取等;另一方面要比较其他常用的实时校正方法,如卡尔曼率波算法,与 AR 模型的 RRLS 算法在有异常误差情况下的效果,验证 RRLS 算法的优越性.

参考文献

[ 1 ] Dupuis DJ, Field CA. Robust estimation of extremes. *Canadian Journal of Statistics*, 1998, 26( 2 ): 199~ 215  
[ 2 ] Huber PJ. Robust estimation of a location parameter. *Ann Math Statist*, 1964, 35: 73~ 100  
[ 3 ] Finsterls S, Najita J. Robust estimation of hydrogeologic model parameters. *Water Resources Research*, 1998, 34( 11 ): 2939~ 2947  
[ 4 ] 周江文, 黄幼才, 杨元喜, 等. 抗差最小二乘法. 武汉: 华中理工大学出版社, 1997. 115~ 116  
[ 5 ] 夏天长. 系统辨识——最小二乘法. 北京: 清华大学出版社, 1983. 40~ 41  
[ 6 ] 周江文. 经典误差理论与抗差估计. *测绘学报*, 1989, 18( 2 ): 115~ 120

Robust recursive algorithm for real-time co-correction model

ZHAO Chao HONG Hua Sheng ZHANG Luo Ping

(State Key Laboratory of Marine Environmental Science, Environmental Science Research Center, Xiamen University, Xiamen 361005, China)

**Abstract** Data observed by telemetric system often carries outliers resulting from instrument malfunctioning, false signal acquisition because of signal leak, collision and disturbance during signal transmission, and special measuring demand of hydrologic variable, in addition to unavoidable random errors. When the parameters of real-time co-correction model are estimated by the abnormal data, the algorithm must not only be able to resist the effect of the outliers, but also have ability for real-time tracing of the changes of parameters. In this paper, a robust recursive least-squares algorithm with a forgetting factor is produced based on the recursive least-squares algorithm. And an example is given to demonstrate that the algorithm is insensitive to the outliers and adapts to the time-varying parameter estimation.

**Key words** co-correction model, robustness, recursive algorithm, forgetting factor